

## BREAST CANCER IMAGE SEGMENTATION USING EKSTRAP AND FCS ALGORITHM

DEEPSHRI HEBBALKAR<sup>1</sup>, SMITA NAIK<sup>2</sup>, ANUJ DIVKAR<sup>3</sup>, AVINKUMAR PEDNEKAR<sup>4</sup>  
VALERIE MENEZES<sup>5</sup> & SHREEDATTA SAWANT<sup>6</sup>

<sup>1,2,3,4</sup>B.E Students, Department of Computer Engineering, Agnel Institute of Technology and Design, Goa, India

<sup>5,6</sup>Assistant Professors, Department of Computer Engineering, Agnel Institute of Technology and Design, Goa, India

### ABSTRACT

*Image segmentation is a process of segmenting the image into different sections. We have done this process of image segmentation using clustering. We have performed EKSTAP and FCS Clustering Algorithms to obtain image segmentation. Then we compared the results with K Means. We noticed that, our technique of obtaining image segmentation gives better results with EKSTRAP and FCS Algorithm. Further, we use a standard classification algorithm called K Nearest Neighbors to identify which section of the image is cancerous and which is non-cancerous.*

**KEYWORDS:** EKSTAP & FCS Clustering Algorithms

**Received:** Apr 14, 2017; **Accepted:** May 15, 2017; **Published:** May 30, 2017; **Paper Id.:** IJCEITRJUN20175

### INTRODUCTION

Breast cancer is the most prominent cancer affecting women. A breast cancer victim's long term survival chances can be improved by prior detection. Since screening mammography is at present the main test for prior detection of breast cancer, a large number of mammograms have to be examined by a limited number of radiologists, resulting in misdiagnosis due to common human errors like visual fatigue [2].

This project provides a second opinion to the radiologist. Hence, using our project the radiologist can give an input image that is a mammogram, and from the output obtained, he/she can accurately distinguish the cancerous regions from the noncancerous regions and hence he/she will be in a better position to identify the cancerous and non-cancerous regions.

Below are some common terms used all through the paper:

**Image segmentation:** Is a process that divides an image into separate and distinct regions.

**Clustering:** A bunch of like entities is a cluster. A method of forming these clusters is known as clustering.

**Classification:** Is a process of sorting sections or segments of images into classes.

### EKSTAP

**EKSTRAP** Algorithm [3] is used to find strange points. The algorithm also dwells on the effect of run time. The EKSTRAP proceeds by initially finding a minimum point by using the Euclidean distance formula [9] for all the pixels of the image. Then it finds the maximum point from the minimum point calculated. The run time

is reduced as it finds the minimum point first and then the maximum point of the minimum point. Next it finds the strange point. If the strange point is closer to the minimum point, then the strange point is rectified by choosing a point which is in the middle of strange point and maximum point and this makes it the final strange point. However, if strange point is closer to a maximum point, then the strange point is shifted from finding a point which is in the middle of the strange point and the minimum point making the newly found out point the final strange point.

## FCS

**FCS** Algorithm [4] takes a broad view of an EKSTRAP Algorithm to permit a point to fit into various clusters by generating a soft categorization for a given set of elements. To achieve this, the motive of hard EKS is drawn out by including fuzzy membership functions. The values of fuzzy membership functions are called memberships range from 0 to 1. Every element has a membership in each cluster. Memberships near 1 would indicate a high degree of association between the element and a cluster. Memberships near 0 implied negligible association between the element and the cluster.

The goal behind using the membership function for clustering was to yield fuzzy, c subdivisions of the dataset. Fuzzy membership functions put the data into the clusters using a parameter  $p$  that has been used as a weight apostrophe in the fuzzy membership function.

## IMAGE SEGMENTATION

In this project, we have segmented the mammogram of a breast tissue, which we have taken as an input image. Hence the image matrix we obtained is shown in the figure below:

**Table 1: Image Matrix of the Input Image**

Pixel Number	210	211	212	213	214	215	216	217
6	0	0	17	17	17	17	17	34
7	17	17	17	17	34	17	34	34
8	17	17	17	17	17	17	34	17
9	17	0	17	34	34	34	34	17
10	17	17	17	34	34	17	34	17
11	34	34	34	34	17	34	17	0
12	34	34	34	34	17	17	0	17
13	34	51	34	17	17	17	17	17

## MOTIVATION

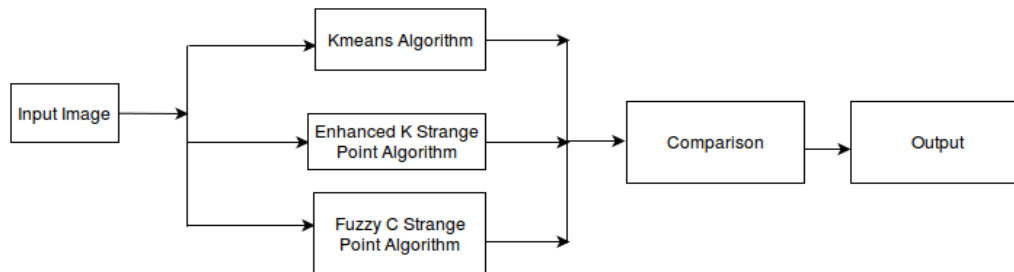
Recently, the numbers of deaths due to breast cancer have increased. A breast cancer victim's survival chances are enhanced by prior detection. Hence, through our project we want to provide a means for early and accurate detection of breast cancer. K Means doesn't comply with huge datasets. Through our research, we found out that K Means is highly time consuming. As the size of the image increases, so does the time required to cluster. It also takes much iteration to reach to the result. EKSTRAP and FCS decrease time complexity. They also undergo a limited number of iterations

## STRATEGY USED

**Image Segmentation:** This is done to divide the image into different sections:

- Take the input image which is a mammogram of a breast tissue.
- Perform clustering using (K Means).

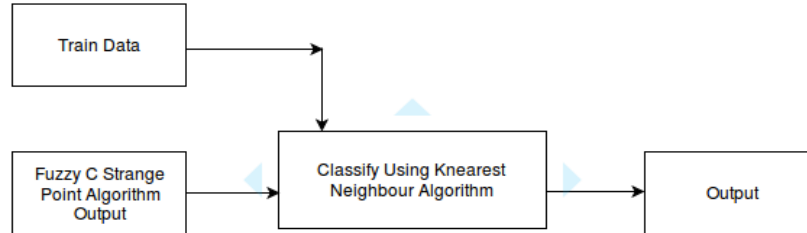
- Perform clustering using (EKSTRAP).
- Perform clustering using (FCS).
- Compare the results obtained from all the 3 Algorithms.
- Produce the output that is a segmented mammogram of a breast tissue.



**Figure 1: Image Segmentation Process**

**Classification:** It is done to assign class labels to the segmented sections of the image

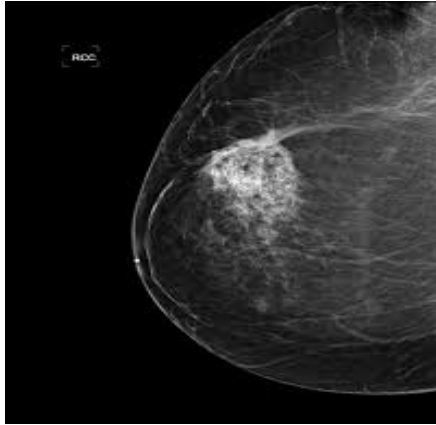
- Train the data using a sample image that is a mammogram of a breast tissue.
- The output from any of the FCS is given to the KNN classifier along with the trained data.
- Final output obtained is an image where the cancerous region is clearly indicated.



**Figure 2: Classification**

## OUTCOME OBTAINED

We first achieve image segmenting using K Means. The steps followed are same as that proposed in the strategy. The input image considered in given below:



**Figure 3: Input Image Selected to Perform K Means**

The clusters obtained from the above image after applying K Means algorithm are shown below:

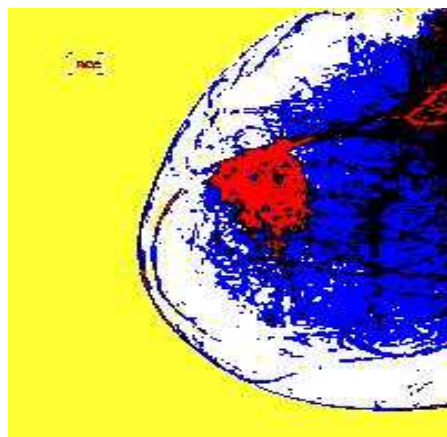
```

Output - BCD_pendrive_1_1_1 (run) x
run:
kmean1: 73.0
kmean2: 65.0
Iteration: 0
Iteration: 1
Iteration: 2
Iteration: 3
Iteration: 4
Iteration: 5
K 0 {397,398,399,400,402,403,599,600,601,800,801,802,803,804,805,806,807,1002,1003,1004,1005,1006,1007,1008,1009,1204,1208,1209,1409,1410,1411,
K 1 {195,196,197,201,350,360,364,365,377,389,390,391,392,393,394,395,396,401,546,547,551,580,588,591,592,593,594,595,596,597,598,602,603,604,60
K 2 {0,1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23,24,25,26,27,28,29,30,31,32,33,34,35,36,37,38,39,40,41,42,43,44,45,46,47,48,4

```

**Figure 4: K Means Forms the Above Clusters**

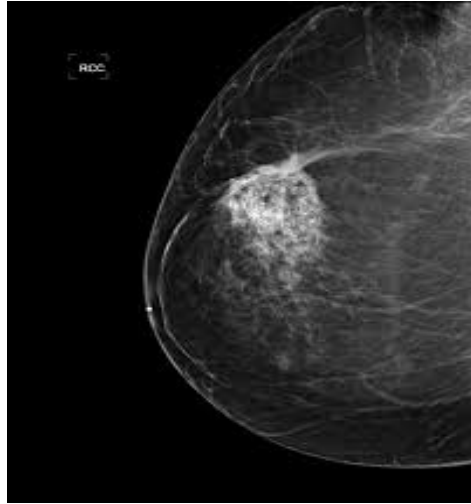
The image obtained after implementing K Means is as shown in the below figure. It segments the image into clusters as shown below. One cluster indicated by the color red, one indicated by the color Blue, one indicated by black, one by white, and another by yellow.



**Figure 5: Output Image after Applying K Means**

## EKSTRAP ALGORITHM

We implement EKSTRAP on the image shown in the figure below, which is taken as the input image:



**Figure 6: Input Image Selected to Perform EKSTRAP**

Clusters formed from the above input image after applying the EKSTRAP are shown in the figure below:

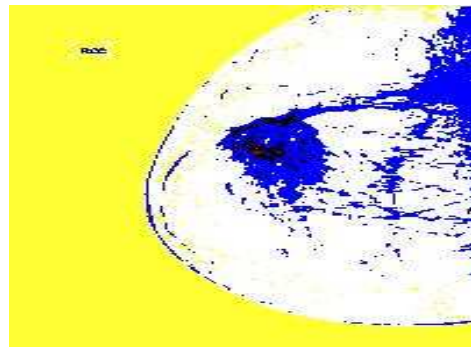
```

Output - BCD_pendrive_1_1_1 (run) x
run:
kmin i 0 j 0: 0
kmax i 105 j 105: 255
The 3 cluster:
Uncorrected strange point: 0.0
Cluster 3 :127.5
cluster1: 217 average: 209
cluster2: 10971 average: 90
cluster3: 39110 average: 16
K 0{6699,16281,16483,16676,16684,16686,16878,16879,16880,16884,16885,16886,16888,17079,17080,17081,17082,17083,17084,17085,17086,17087,17088,1
K 1{196,395,396,397,398,399,400,401,402,403,547,597,598,599,600,601,602,603,604,605,745,746,752,797,798,799,800,801,802,803,804,805,806,807,94
K 2{0,1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23,24,25,26,27,28,29,30,31,32,33,34,35,36,37,38,39,40,41,42,43,44,45,46,47,48,4

```

**Figure 7: EKSTRAP Forms the Above Clusters**

The output image obtained after applying EKSTRAP is as shown in the below figure. It segments the image into clusters as shown below. One cluster indicated by the color red, one indicated by the color Blue, one indicated by black, one by white, and another by yellow. The output obtained precisely segments the image.



**Figure 8: Output Image after Applying EKSTRAP**

## FCS ALGORITHM

This algorithm segments the image better than the K means. It also runs faster than the K Means when a large dataset is considered. We implement FCS on the image shown in the figure below that is taken as the input image:

Clusters formed from the above input image after applying FCS are shown in the figure below:

```

Output - BCD_pendrive_1_1_1(run) X
run:
fmin i 0 j 0: 0
fmax i 105 j 105: 255
The 3 cluster:
Uncorrected strange point: 0.0
Cluster 3 :127.5
After sorting:
255.0
127.5
0.0
cluster1: 217 average: 208.8847926267281
cluster2: 10971 average: 90.02360769300884
cluster3: 39110 average: 16.169521861416516
K 0{6699,16281,16483,16676,16684,16686,16878,16879,16880,16884,16885,16886,16888,17079,17080,17081,17082,17083,17084,17085,17086,17087,17088,1
K 1{196,395,396,397,398,399,400,401,402,403,547,597,598,599,600,601,602,603,604,605,745,746,752,797,798,799,800,801,802,803,804,805,806,807,94
K 2{0,1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23,24,25,26,27,28,29,30,31,32,33,34,35,36,37,38,39,40,41,42,43,44,45,46,47,48,4

```

Figure 9: Input Image Selected to Perform FCS

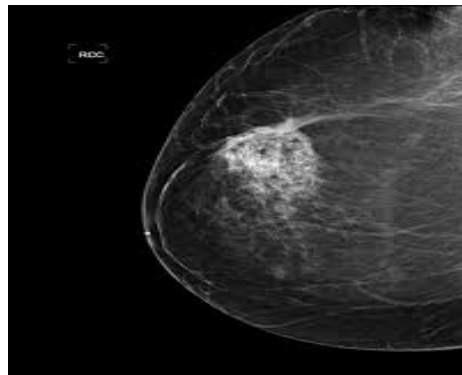


Figure 10: FCS Forms the Above Clusters

The output image obtained after applying FCS is as shown in the below figure. It segments the image into clusters as shown below. One cluster indicated by the color red, one indicated by the color Blue, one indicated by black, one by white, and another by yellow. The output obtained precisely segments the image.

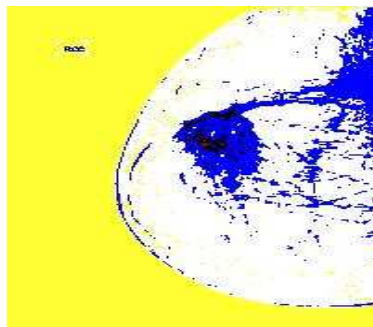
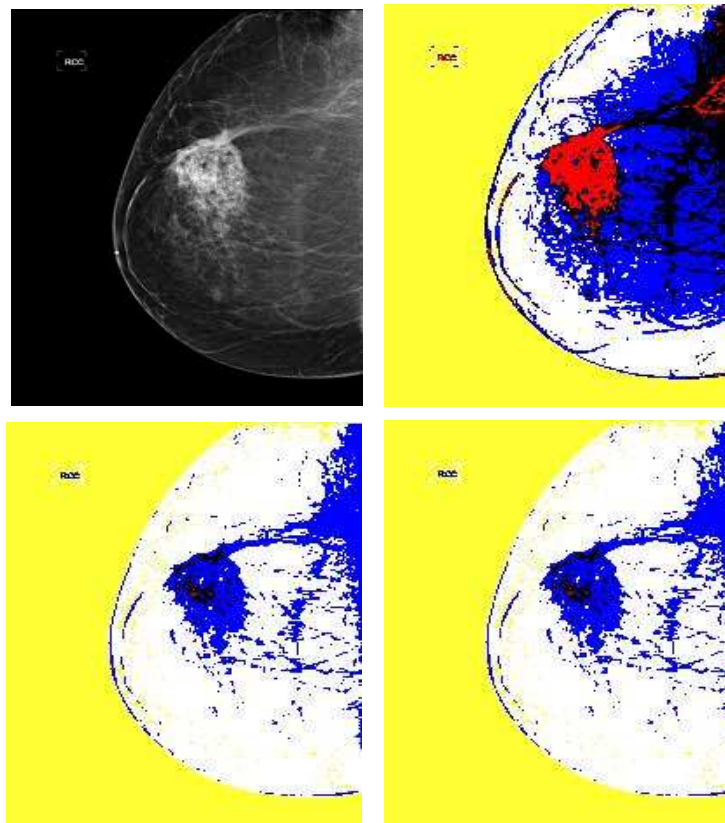


Figure 11: Output Image after Applying FCS Clustering

### Difference with Respect to K Means



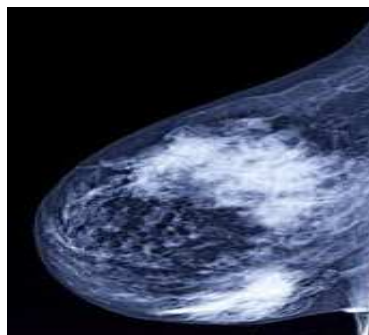
**Figure 12: (A) the Input Image. (B) Output of K Means. (C) Output of EKSTRAP. (D) Output of FCS**

The above figure clearly shows how precisely and accurately the EKSTRAP and the FCS works as compared to the K Means Clustering algorithm.

### KNN CLASSIFIER

We have implemented the KNN Classifier [5] to allocate labels to the segmented sections of the image. We followed the similar steps as shown in the proposed methodology under classification.

We train the data using a sample image as given in the figure below:



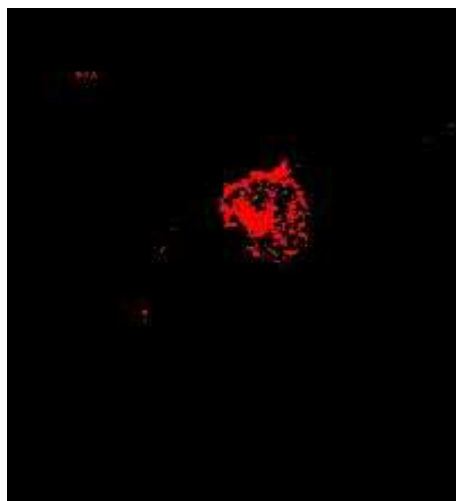
**Figure 13: Sample Image Taken to Train the Data**

We train the data by setting the cancerous region to 1 and the noncancerous region to 0. In other words we train the image pixels according to their intensity as cancerous and noncancerous. This is shown in the below figure:

**Table 2: The Trained Data as 1(Cancerous) and 0 (Noncancerous)**

	A	B
1	162.063	1
2	29.38539	0
3	208.8848	1
4	90.02361	0
5	16.16952	0
6	211.1579	1
7	154.4296	0
8	80.29887	0
9	14.9067	0
10	245.0909	1

The output obtained from training the data along with the output obtained from image segmentation using FCS is given to the KNN classifier. The KNN Classifier shows which regions are cancerous and which are not. The following figure highlights the cancerous region.



**Figure 14: Output of KNN Classifier**

In the above figure, the region marked in red is the cancerous region, whereas the one marked in black is the noncancerous region.



## CONCLUSIONS

In this paper, we have obtained image segmentation. We have used EKSTRAP and FCS to obtain image segmentation. The results obtained from these two algorithms were compared with the results obtained from K Means. Our technique of obtaining image segmentation gives better results with EKSTRAP and FCS Algorithm. We then classified the segmented image using a KNN classifier, to indicate which sections of the image are cancerous and which are noncancerous

## REFERENCES

1. Tntu P B, Paulin.R(2013) "Detect Breast Cancer Using Fuzzy C-Means Techniques In Wisconsin Prognostic Breast Cancer(WPBC) Datasets" *International Journal Of Computer Applications And Technology And Research*
2. Nalini Singh, Ambarish G. Mohapatra, Gurukalyan Kanungo (2011) "Breast Cancer Mass Detection In Mammograms Using K-means and Fuzzy C-means Clustering" *International Journal of Computer Applications* (0975-8887)
3. Terence Johnson, Dr. Santosh Kumar Singh,(2015) "Enhanced K Strange Points Clustering Algorithm Proceedings of the 2nd International Research Conference on Emerging Information Technology and Engineering Solutions". EITED 2015, 978-1-4799-1838/15, IEEE Computer Society Washington, DC USA 2015 IEEE, DOI 10.1109/EITS.2015.14, indexed in ACM Digital Library, pp. 32-37
4. Terence Johnson, Santosh Kumar Singh,(2016) "Fuzzy C Strange Points Clustering Algorithm" *IEEE International Conference On Information Communication And Embedded Systems(ICICES 2016)* "ICICES 2016", February 25-26,2016 S A Engg College, Chennai, India. E-ISBN-978-1-5090-2552-7, PoD ISBN: 978-1-5090-2553-4, INSPEC Accession Number:161596528, DOI:10.1109/ICICES.2016.751829, IEEE, pp1-5
5. Tsung-Hsien Chiang, Hung-Yi Lo, Shou-De Lin, "A Ranking-based KNN Approach for Multi-Label Classification", *JMLR: Workshop and Conference Proceedings* 25:81–96, 2012 Asian Conference on Machine Learning
6. S. Saheb Basha, Dr. K. Satya Prasad(2009) "Automatic Detection of Breast Cancer Mass In Mammograms Using Morphological Operators and Fuzzy C-means Clustering" *Journal of Theoretical and Applied Information Technology*
7. Jiawei Han and Micheline Kamber, *Data Mining- Concepts and techniques*, Elsevier.
8. Terence Johnson, Dr. Santosh Kumar Singh, Valerie Menezes, Edrich Rocha, ShriyanWalke, DikshaPrabhuKhorjuvekar, Sana Pathan, "Improved Steganography using Enhanced K Strange Points Clustering" *International Journal of Applied Engineering Research* ISSN 0973-4562 Volume 11, Number 9 (2016) pp 6881-6885 © Research India Publications. <http://www.ripublication.com>
9. A. Alfakih, A. Khandani and H. Wolkowicz, "Solving Euclidean Distance Matrix Completion Problems", *Comput. Optim. Appl.* 12(1999), pp. 13-30.
10. Image database: <http://peipa.essex.ac.uk/info/mias.html>
11. Image data [https://www.researchgate.net/post/Free\\_database\\_for\\_breast\\_cancer\\_biopsy\\_images](https://www.researchgate.net/post/Free_database_for_breast_cancer_biopsy_images)

